

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 30, 2002, 17:44:56 ; Search time 115.69 Seconds  
(without alignments)  
1235.267 Million cell updates/sec

Title: US-09-357-273a-2  
Perfect score: 5139  
Sequence: 1 MPARLLLLTLPLGLGIF.....QPYFHPPEPPQPVTPDAL 977

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Archived: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_protent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5139	100.0	977	4	075460
2	4741	92.3	977	11	09EOYO
3	2192.5	42.7	911	11	0922E3
4	2017	39.2	408	11	09D340
5	1604.5	31.2	1038	5	09VDT7
6	959.5	18.7	1072	3	094537
7	770	15.0	893	10	09FS17
8	755.5	14.7	939	10	09FIN6
9	727	14.1	841	10	09CSM2
10	655.5	12.8	393	10	09SHL6
11	529.5	10.3	354	10	09SF12
12	416	8.1	169	1	09C4S9
13	334	6.5	1114	11	0922B5
14	320	6.2	1230	3	09HFT9
15	312	6.1	1108	11	09Z1Z1
16	309	6.0	666	3	042625
17	306	6.0	735	11	09ERU7
18	299.5	5.8	536	3	09P466
19	297.5	5.8	1077	5	019192

20	295	5.7	1257	10	064768	064768 arabidopsis
21	293.5	5.7	360	5	094608	094608 limulus pol
22	290	5.6	527	13	09PW62	09PW62 xenopus lae
23	290	5.6	793	3	09HRW2	09HRW2 ashbya goss
24	286.5	5.6	360	5	094609	094609 limulus pol
25	286	5.6	842	10	09FNU3	09FNU3 oryza sativ
26	285	5.5	311	5	096526	096526 leishmania
27	285	5.5	525	13	057318	057318 xenopus lae
28	283.5	5.5	530	3	042793	042793 colletoic
29	283.5	5.5	823	3	09HFR3	09HFR3 pneumocysti
30	283.5	5.5	1152	5	017346	017346 caenorhabdi
31	283	5.5	311	5	015851	015851 leishmania
32	283	5.5	350	5	09NG91	09NG91 caenorhabdi
33	283	5.5	482	5	09U600	09U600 caenorhabdi
34	283	5.5	518	5	09NH57	09NH57 caenorhabdi
35	283	5.5	520	5	09NH60	09NH60 caenorhabdi
36	282	5.5	302	5	09NH56	09NH56 caenorhabdi
37	282	5.5	302	5	09NH55	09NH55 caenorhabdi
38	282	5.5	533	5	09NH59	09NH59 caenorhabdi
39	282	5.5	571	5	09NH58	09NH58 caenorhabdi
40	281	5.5	547	5	021431	021431 caenorhabdi
41	280.5	5.5	1271	10	09LX30	09LX30 arabidopsis
42	279	5.4	493	5	09V495	09V495 drosophila
43	279	5.4	825	10	09AKL7	09AKL7 hordeum vul
44	279	5.4	1054	10	09FKL3	09FKL3 arabidopsis
45	279	5.4	1096	5	017368	017368 caenorhabdi

## ALIGNMENTS

RESULT 1  
ID 075460 PRELIMINARY; PRT; 977 AA.  
AC 075460;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE PROTEIN KINASE/ENDORIBONUCLEASE.  
GN IREL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_Taxid:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA MEDLINE=98301437; PubMed=9637683;  
RX Tirasophon W., Welihinda A.A., Kaufman R.J.;  
RT "A stress response pathway from the endoplasmic reticulum to the  
RT nucleus requires a novel bifunctional protein kinase/endoribonuclease  
RT (irelp) in mammalian cells."  
RL Genes Dev. 12:1812-1824 (1998).  
CC EMBL: AF059198; AAC25991.1;  
DR InterPro: IPR002372; Bac\_P00.repeat.  
DR InterPro: IPR000719; Euk\_Pkinase.  
DR InterPro: IPR001005; Myb\_DNA\_bind.  
DR InterPro: IPR002290; Ser\_Thr\_kin\_actsite.  
DR Pfam: PF00069; Bacterial\_P00; 2.  
DR Pfam: PF00069; Bacterial\_P00; 2.  
DR PROSITE: PS00037; MYB\_1; UNKNOWN.1.  
DR PROSITE: PS00037; MYB\_1; UNKNOWN.1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_DOM.1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST.1.  
KW ATP-Binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 977 AA; 109686 MW; EAFSEI082D09010C CRC64;

Query Match 100.0%; Score 5139; DB 4; Length 977;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 977; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
QY 1 MPARLLLLTLPLGLGIFGSTSTVTLPERLLFVSTLDGSLAAVSKRTGSIKWTLEKDP 60

```

Db      1 MPAKRLLLLLLLGLGIFGTSSTVTLPELTLFVSTLDGSLHAASKRTGSIKMTLKEDEP 60
Qy      61 VLOVPTHEEPAPLPDPNDGSLYTLGSKNNNGTLPTPELTVASACRSSDGLTYMGK 120
Db      61 VLOVPTHEEPAPLPDPNDGSLYTLGSKNNNGTLPTPELTVASACRSSDGLTYMGK 120
Qy      121 KODIMVYIDLTLGKCOQTLSAFAADSLCPSTSLYLGRTREYITMYDPTKRELRNATYF 180
Db      121 KODIMVYIDLTLGKCOQTLSAFAADSLCPSTSLYLGRTREYITMYDPTKRELRNATYF 180
Qy      121 KODIMVYIDLTLGKCOQTLSAFAADSLCPSTSLYLGRTREYITMYDPTKRELRNATYF 180
Db      121 KODIMVYIDLTLGKCOQTLSAFAADSLCPSTSLYLGRTREYITMYDPTKRELRNATYF 180
Qy      181 DYASLPDEBDYKMSHFVNSGDCLVTVVDSGCVLIQNTASPVAFYVWQREGLRKY 240
Db      181 DYASLPDEBDYKMSHFVNSGDCLVTVVDSGCVLIQNTASPVAFYVWQREGLRKY 240
Qy      241 MHINAVETLRYLPFMSGEVGRITKMKYPRPKETBAKSLPTLVGAYSTSLYASPSMV 300
Db      241 MHINAVETLRYLPFMSGEVGRITKMKYPRPKETBAKSLPTLVGAYSTSLYASPSMV 300
Qy      301 HEGVAVVPRGSTLPLEGGPQTDGVTIGDKGECVITPSTDVKEDPLGSKNKLNYLRNWL 360
Db      301 HEGVAVVPRGSTLPLEGGPQTDGVTIGDKGECVITPSTDVKEDPLGSKNKLNYLRNWL 360
Qy      361 LIGHETPLASTKMLEFPNNLPKHRENVIPADSEKKSFEVITNLVQTSNATYTSR 420
Db      361 LIGHETPLASTKMLEFPNNLPKHRENVIPADSEKKSFEVITNLVQTSNATYTSR 420
Qy      421 DVEEKPAHAPRPAVPSMLKDMATILSTFLIGWAFITTYPLSMHQOOLQHOQFO 480
Db      421 DVEEKPAHAPRPAVPSMLKDMATILSTFLIGWAFITTYPLSMHQOOLQHOQFO 480
Qy      481 KELEKIQLLQOQOOLPFPHPGDTAODGELLDTSGPYSESSGTSPTSPPRASNSHLSG 540
Db      481 KELEKIQLLQOQOOLPFPHPGDTAODGELLDTSGPYSESSGTSPTSPPRASNSHLSG 540
Qy      541 SSASKAGSSPLEDDDGDETSVYVKGISFCPKVYLGHGAGTIVYGMGMDNRVAAKR 600
Db      541 SSASKAGSSPLEDDDGDETSVYVKGISFCPKVYLGHGAGTIVYGMGMDNRVAAKR 600
Qy      601 ILPECFSPADREVOYLRSDEHPNVIYFCTEKRQFOYIAELCAATLOEYVEQKDFAH 660
Db      601 ILPECFSPADREVOYLRSDEHPNVIYFCTEKRQFOYIAELCAATLOEYVEQKDFAH 660
Qy      661 LGLEPTLLOOTSGLAHLSLNYHDLKPHNLIISMPNAHGKIKAMISDFGLCKLAV 720
Db      661 LGLEPTLLOOTSGLAHLSLNYHDLKPHNLIISMPNAHGKIKAMISDFGLCKLAV 720
Qy      721 GRHSFSRSGVPTGEGWIAPEMLSDCKENPTYVDIFSACVFEYVVSSEGSHPRGKSLQ 780
Db      721 GRHSFSRSGVPTGEGWIAPEMLSDCKENPTYVDIFSACVFEYVVSSEGSHPRGKSLQ 780
Qy      781 ROANILLGACSLDCLHPKHEHDIARELIEKMIAMDPOKRSANDVILKHPFWSLEKOLQ 840
Db      781 ROANILLGACSLDCLHPKHEHDIARELIEKMIAMDPOKRSANDVILKHPFWSLEKOLQ 840
Qy      841 FFQOVSDRIEKSLEDGPIVQOLERGRAVVKMDRENITDPLQTDLRKFRYKGGSVBDL 900
Db      841 FFQOVSDRIEKSLEDGPIVQOLERGRAVVKMDRENITDPLQTDLRKFRYKGGSVBDL 900
Qy      901 LRAHRNKNHNRRELPAVEHRETLGLTPDQFVYFYSRPHLANTYRAMELCSHELLRPY 960
Db      901 LRAHRNKNHNRRELPAVEHRETLGLTPDQFVYFYSRPHLANTYRAMELCSHELLRPY 960
Qy      961 YFHERPERQPVTPDAL 977
Db      961 YFHERPERQPVTPDAL 977

```

```

DT      01-MAR-2001 (Tremblrel. 16, last sequence update)
DE      01-JUN-2001 (Tremblrel. 17, last annotation update)
Db      PROTEIN KINASE/ENDORIBONUCLEASE (IRE1) ALPHA.
GN      IRE1 OR IRE1 ALPHA
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      PubMed=11146108;
RA      Miyoshi K., Katsuyama T., Imaiuchi K., Taniguchi M., Mori Y.,
RA      Hitomi J., Yui D., Manabe T., Gomi F., Yoneda T., Tohyama M.;
RT      "Characterization of mouse Irelalpha: cloning, mRNA localization in
RT      the brain and functional analysis in a neural cell line.";
RL      Mol. Brain Res. 85:68-76(2000).
DR      EMBL: AB031332; BAB20901.1; -
DR      MGI:1930134; Irela.
DR      InterPro: IPR002372; Bac_PQO_repeat.
DR      InterPro: IPR000719; Euk_Pkinase.
DR      InterPro: IPR002290; Ser_thr_kin_actsite.
DR      pfam: IPR001245; Tyr_kin.
DR      pfam: PF01011; Bacterial_PQO_2.
DR      pfam: PF00063; Pkinase_1.
DR      SMART: SM00220; S_TKC_1.
DR      SMART: SM00219; TyKc_1.
DR      PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR      ATP-binding; Kinase; Transferase.
SQ      SEQUENCE 977 AA; 110184 MW; 216E3E2PA2FF370 CRC64;

Query Match      92.3%; Score 4741; DB 11; Length 977;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 903; Conservative 34; Mismatches 38; Indels 4; Gaps 2;

Qy      1 MPAKRLLLLLLL--PGLGIFGTSSTVTLPELTLFVSTLDGSLHAASKRTGSIKMTLKE 58
Db      1 MPAKRLLLLLLLLPPEPGSFGRTSTVTLPELTLFVSTLDGSLHAASKRTGSIKMTLKE 60
Qy      59 DPLVOPPTHEEPAPLPDPNDGSLYTLGSKNNNGTLPTPELTVASACRSSDGLTYMGK 118
Db      61 DPLVOPPTHEEPAPLPDPNDGSLYTLGSKNNNGTLPTPELTVASACRSSDGLTYMGK 120
Qy      119 GKODIMVYIDLTLGKCOQTLSAFAADSLCPSTSLYLGRTREYITMYDPTKRELRNAT 178
Db      121 GKODIMVYIDLTLGKCOQTLSAFAADSLCPSTSLYLGRTREYITMYDPTKRELRNAT 180
Qy      179 YEDFAASLPDEBDYKMSHFVNSGDCLVTVVDSGCVLIQNTASPVAFYVWQREGLR 238
Db      181 YEDFAASLPDEBDYKMSHFVNSGDCLVTVVDSGCVLIQNTASPVAFYVWQREGLR 240
Qy      239 KVMHINAVETLRYLPFMSGEVGRITKMKYPRPKETBAKSLPTLVGAYSTSLYASPS 298
Db      241 KVMHINAVETLRYLPFMSGEVGRITKMKYPRPKETBAKSLPTLVGAYSTSLYASPS 300
Qy      299 MHESGAVVPRGSTLPLEGGPQTDGVTIGDKGECVITPSTDVKEDPLGSKNKLNYLRN 358
Db      301 MHESGAVVPRGSTLPLEGGPQTDGVTIGDKGECVITPSTDVKEDPLGSKNKLNYLRN 360
Qy      359 WLLGHNHTPLASATKMLEFPNNLPKHRENVIPADSEKKSFEVITNLVQTSNATYTSR 418
Db      361 WLLGHNHTPLASATKMLEFPNNLPKHRENVIPADSEKKSFEVITNLVQTSNATYTSR 420
Qy      419 SRDVEKPAHAPRPAVPSMLKDMATILSTFLIGWAFITTYPLSMHQOOLQHOQFO 478
Db      421 SRDVEKPAHAPRPAVPSMLKDMATILSTFLIGWAFITTYPLSMHQOOLQHOQFO 480
Qy      479 FOKLEKIQLLQOQOOLPFPHPGDTAODGELLDTSGPYSESSGTSPTSPPRASNSHLSG 538
Db      481 FOKLEKIQLL--QOQOOLPFPHPGDTAODGELLDTSGPYSESSGTSPTSPPRASNSHLSG 538
Qy      539 SGSSKAGSSPLEDDDGDETSVYVKGISFCPKVYLGHGAGTIVYGMGMDNRVAAKR 598
Db      539 SGSSKAGSSPLEDDDGDETSVYVKGISFCPKVYLGHGAGTIVYGMGMDNRVAAKR 598

```

RESULT 2  
 Q9EQYO 09EQYO PRELIMINARY; PRT: 977 AA.  
 AC 09EQYO: 01-MAR-2001 (Tremblrel. 16, Created)

Dh 539 PSSASRAGISPSLEODEDEETRMVIVKISPCPKDVLHGAGCTIYVKGMFNDNDVAV 598  
QY 599 KRILPECSFADREVQLRESDEHPNVIYRFTCEKROFOYIAIETCAILOEYEQKF 658  
Dh 599 KRILPECSFADREVQLRESDEHPNVIYRFTCEKROFOYIAIETCAILOEYEQKF 658  
QY 659 AHLGLEPTLLQOTTSGLAHLHSLNIVHRLDKPHNIIISMPNAGKIKAMISDFGLCKL 718  
Dh 659 AHLGLEPTLLQOTTSGLAHLHSLNIVHRLDKPHNIIISMPNAGKIKAMISDFGLCKL 718  
QY 719 AVGRHSFRSGVPTCEGWIAPEMISEDCNDPTTYVDFISAGCYFYVVSQSPFGKS 778  
Dh 719 AVGRHSFRSGVPTCEGWIAPEMISEDCNDPTTYVDFISAGCYFYVVSQSPFGKS 778  
QY 779 LORQANILGACSLDCLPEKHEDVIARELLEKMIAMPORPSANDVLKHPFSLERQ 838  
Dh 779 LORQANILGACSLDCLPEKHEDVIARELLEKMIAMPORPSANDVLKHPFSLERQ 838  
QY 839 LOFQODVSDRIEKESLDSPIVKQLEGRGAVVKKMDRENITDPLQDLRKFRYKGSVR 898  
Dh 839 LOFQODVSDRIEKESLDSPIVKQLEGRGAVVKKMDRENITDPLQDLRKFRYKGSVR 898  
QY 899 DLLRAMRKKHHYRLPVEVRETGLTLPDDVYCYTTSRPHILANTYAMELCSHERLFQ 958  
Dh 899 DLLRAMRKKHHYRLPVEVRETGLTLPDDVYCYTTSRPHILANTYAMELCSHERLFQ 958  
QY 959 PYFHEPEPOPVPYPPAL 977  
Dh 959 PYFHEPEPOPVPYPPAL 977  
RESULT 3  
ID 0922E3 PRELIMINARY: PRT: 911 AA.  
AC 0922E3  
DT 01-MAY-1999 (TREMblrel. 10, Created)  
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE IREL.  
GN IREL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
SEQUENCE FROM N.A.  
MEDLINE=98429494; PubMed=9755171;  
Wang X.Z., Harding H.P., Zhang Y., Jolicœur E.M., Kuroda M., Ron D.;  
"Cloning of mammalian Irel reveals diversity in the ER stress  
responses.";  
EMBO J. 17:5708-5717(1998).  
CC -I- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL: AF071777; AAC64400.1; -.  
DR MGI: MGI:1349436; Irel.  
DR InterPro: IPR002372; Bac\_PQO\_repeat.  
DR InterPro: IPR000719; Euk\_kinase.  
DR InterPro: IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam: PF01011; Bacterial\_PQO; 1.  
DR Pfam: PF00069; Kinase; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS50108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SEQUENCE 911 AA; 101355 MW; 173EDE163841FE7 CRC64;

Query Match 42.7%; Score 2192.5; DB 11; Length 911;  
Best Local Similarity 47.7%; Pred. No. 2.8e-159;  
Matches 478; Conservative 123; Mismatches 267; Indels 135; Gaps 20;

QY 2 PARRLILLTLLPLGLFGSTSTVTLPEPTLFGSTLGSGLAVKRRGSIKWTLEKDPV 61  
Dh 13 PLGLLQVLTLL-----LGLKLPQVQSVRPESLTFVSTWDSGLHALNKQTEGLKWTVKDDPI 68

QY 62 LQVPHVEEPAFLPDPNDGSLYTLGSKNNEGITLKLPTIPELVQASPCRRSDGILYMGKK 121  
Dh 69 IGGPMYVTEMAFLSPADGASLYVLGTHKLOGIMKLPFTIPELVHASPCRRSDGVEYTGK 128  
QY 122 QDIWVIDLLTGEKQOTSSAFAADSLCSTSLIYIGREFYITMTDTRELRNATYFD 181  
Dh 129 QDAFVVPDESGETMTLT---EGL--STPQLFGRQYVSMHDLRTPALRMNTYR 183  
QY 182 YAASLPDEGDYKMSHFNSNGDLYVTVDESQDVLMTQNVASPVAVVQOREGLARKM 241  
Dh 184 YSAPLINSFGKYMSHLISCGMGLITVDPGSGLYLMQDGLGVPVGTGYTHWQDLHQLP 243  
QY 242 HINVAEVLRLYLTFNSGEGVRLTKWKYPKETEAK---SKLPITYLGYKYSTLYAS 296  
Dh 244 HLTARDLTLHLFLRMGHI--RLPASSY---OPTAQFSSLDLQMLTYLVGKEEAGFYVS 299  
QY 297 PSMVEGVAANVRGSLPLLEGPQTDGVTIGKGEVITPSTVDYKFPDGLSKNKLNTLR 356  
Dh 300 KALVAGVALVYRGLTLPMDGPTIDEVTLQVSGERESPSTAVRY PGSSVA-----LP 353  
QY 357 NYWLLIGHETPLASTKMLERFPNNLPKHRENVLPADSEKKS-----FEEVINYDQ 409  
Dh 354 SQMLLIGHETPPVYLHTMLRVHP--IP-----GKYSAEFRASEDLHARVPFELLNRE 407  
QY 410 TSENAPTYVSRDVEKPAHAPAPAPVDSM-----LKDMATILSTFLLIGVAFILTYR 465  
Dh 408 DEPLRP-----EEKS-----DSYGLGSQDLAATFPAILLGAAVLY----- 445  
QY 466 LSMHQQOOLQHQOFQEKLEKIOLOOQOQOLPFPH--REDTAQDG-----ELDTSGPY 517  
Dh 446 -----LMROOQSPSAPAGPDLQDAQOLSRDILDOORF 482  
QY 518 SESSGTSPTSPPRASNSHSLCGSSASAKASSPSLEQDDGDETSVIVKISPCPKDL 577  
Dh 483 QSPSEPAQPPHDEAGQPT-----VVGKISFNPKDL 514  
QY 578 GHAGAGTYIYRGMPFNRDVAVKRLPECFSPADREVQLRESDEHPNVIYRFTCEKROF 637  
Dh 515 GRAGAGTYVFRQFEGRAVAVKRLRECGYLRREVQLQESDRHPNVIYRFTCEHGFQF 574  
QY 638 QYIAELCATILOEYEQKDFANHLGLEPTLLQOTTSGLAHLHSLNIVHRLDKPHNIIIS 697  
Dh 575 HYIAELCAISOEYVESDLDRLWGLEPTTYVLOOMMSGLAHLHSLNIVHRLDKPHNIIIS 634  
QY 698 MPNAGKIKAMISDFGLCKLAVGRHSFRSGVPTCEGWIAPEMISEDCXCKNPTTYDI 757  
Dh 635 GPDSSQGGKVVYLDPLGCKLPVGRCSFSLHSGIPETBSMMAPELLOLP--PDSPTNAVDI 693  
QY 758 FSAGCVFYVVSSESHPEKSLQROANILGACSLDCLPEKHEDVIARELLEKMIAMP 817  
Dh 694 FSAGCVFYVVSSESHPEKSLQROANILGACSLDCLPEKHEDVIARELLEKMIAMP 753  
QY 818 QKRPASANDVLKHPFWSLEKQLOFQODVSDRIEKESLDSPIVKQLEGRGAVVKKMDREN 877  
Dh 754 QDRPSAGWLAHPLFWMSRAKELQFQODVSDRIEKESLDSPIVKQLEGRGAVVKKMDREN 813  
QY 878 TNDPLQDLRKFRYKGSVGRDLRLAMRKKHHYRLPVEVRETGLTLPDDVYCYTTSR 937  
Dh 814 ISAPLQADLKRSTYKGSVGRDLRLAMRKKHHYRLPVEVRETGLTLPDDVYCYTTSR 873  
QY 938 PHLAHTYAMELCSHERLFQPYFHEPEPOPVPY---TPDA 976  
Dh 874 PRLLLHTHRAMRTCASSESLFLPY-----PRLAERARPA 908

RESULT 4  
ID 09D340 PRELIMINARY: PRT: 408 AA.  
AC 09D340;  
DT 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE 9030414B18RIK PROTEIN.

	Query Match	Similarity	95.8%	Score	2017:	DB 11:	Length	408:	
	Best Local	Similarity	95.8%	Pred.	No. 2_2e-146:				
	Matches	388:	Conservative	5:	Mismatches	10:	Indels	2:	Gaps
Qy	1	MPARLLLLLL--PGLGIFGSTVYTLPELTVSTLDGSLHAVSKRTGISIKWTLKE	58						
Dd	1	MPARWLLELLLPPPGSGSFGRTSVTLPETLLFVSTLDGSLHAVSKRTGISIKWTLKE	60						
Qy	59	DPVLVPFHVEEPALPDNDGSLVTIGSKNNEGLTKLPFTIPELVOASPCRSSDILM	118						
Dd	61	DPVLVPFHVEEPALPDNDGSLVTIGKKNNEGKLKPLFTIPELVOASPCRSSDILM	120						
	119	GKKODIWVIDLTGEKOOTLSAFADSLCPSTSLYLGRTEYTYMYDJTKRELWMNT	178						
	121	GKKODIWVIDLTGEKOOTLSAFADSLCPSLSLYLGRTETTYMYDJTKRELWMANT	180						
Qy	179	YEDVAASLPDEGDGXMSHFNVSNGGLVYTVSESGDVLIOMNVASPVAAFYVMQREGR	238						
Dd	181	YFDVAASLPDEDDVDYDKMSHFVNSGGGLVYTVSESGDVLIOMNVASPVAAFYVMQDEIVR	240						
Qy	239	KVMHINVAEFLRLRYLTFMSGEGVRTKKMKPRPKETEAKSKLPTLYVGKYSTSLASP	298						
Dd	241	KVMHINVAEFLRLRYLTFMSGEGVRTKKMKPRPKETEAKSKLPTLYVGKYSTSLASP	300						
Qy	299	MVHEVAAVPRGCSLPLLEGQTQDGVITGDKECVITPSTDVVKFPDGLKSNNKLANLRNY	358						
Dd	301	MVHEVAAVPRGSTPLPLEGPQTQDGVITGDKCECVITPSTDVVKFPDGLKGSKNLANLRNY	360						
Qy	359	WLLIGHNETPLASTKMLERFPNNLPKRHENYIPADSEKSKFEVY 403							
Dd	361	WLLIGHNETPLASTKMLERFPNNLPKRHENYIPADSEKRSFEVY 405							
RESULT	5								
99VDT7									

```
Query Match      31.2%   Score 1604.5; DB 5; Length 1038;
Best Local Similarity 37.5%   Pred. No. 4.6e-114;
Matches 379; Conservative 152; Mismatches 276; Indels 203; Gaps 31.
```

```
Db 27 NVQVPHFLPDRDGSIVYLGQWGS--LKKLPYIPOLVANAPCRSSDGILYSGKSDTWY 84
OY 127 VIDLITGKQOQLSSAFADSLC-----PSTSLYLGRTETITITVDF--KTRELR 174
Db 85 MYDPTGKREKVM--GFDATVVGKEGQIGWATSRATYLGRTQYVMYDLSLAKKDK 142
OY 175 -WNAIYFYAA--SLPEDEGDYKMSHFVNSGDLVYTVDSGVDLMIQNYASPVAAFYV 231
Db 143 PNIITFYDYNVNSAPPELAKKEYEYIHLTTTNGQIVTIDRKIKGKFLMODLSSPVAAFL 202
OY 222 WOREGLRKVMHINAVETLRVLTWMS--GEVGRITKMKYPPPKETEAASKLPTLYVGY 289
Db 203 LGPDDLSPFTTVDQVDEAVQALLESKKTGNVNTV-----KLFQSLVGEH 247
OY 290 SLSLQASMSVMEGVAVVPDSTLP---LLBGPQDGVYIGDKGECVITPSPDVKFPDL 346
Db 248 QKGLALPSLVAKN--TPRISTAPPIKLDSP-----TGQNSNGEEDPRTIYINDVL 298
OY 347 KSKNKLNTLRNWLIGHHE-----TPLSASTKMLRFRPNNLPKHREN---VIP 392
Db 299 Q-----EHAGIMLGHTNMPNEGNGNLQSLPTASSKVVQSLAT---IHNYNDGYGL 347
OY 303 ADSEKKSFE-----EYINL-VDO---TSENAPTVSRDVEEKPAHA---PARPA 435
Db 348 ANNEKNADIGVQTDPELVEIGDQRTNGNTINRTKITLQNSNKKVQAFINEMFMEHPSG 407
OY 436 PYDSMLKQKATITLSTFLLIGVAVITITYPLSMHQOOLQHQFOKELEKIQLLDQOQOQ 495
Db 408 KWHQIL-----IVYLGMI-A-LFWYTCST-----MKELOK----- 436
OY 496 LPFHPGDTAODELLDTSGPYSESSGTSPTSPRASNSHLSGSSAKAGSPSELOD 555
Db 437 -----QSENGSKTFALQNGSNGS---TSGNSGNANAEDELVDIG 472
OY 556 DQDEETSVYVIGKISFCPKVDLGHAEGETIVYRGMPDNDAVVKRILPCFSFADREVOL 615
Db 473 NQO-----VRGKISFSTNEVVGKCEGTFVFGTEEFVFKRLLPCEFTFADREVAL 527
OY 616 LRESDEHPVIVIFCTEKKOROFQYAIETLCATLOEYVQKQFANL--GLEPITLIQOT 673
Db 528 LRESDAHEKVVYKFCEDQKQFRTAVELCAATLQDYTEGDRSLEQNHIDWQVLSQAA 587
OY 674 SGLAHLHSINIVHDLKPHNIIISMNPAHGKIKAMISDFGLCKLAVGHSEFSRSRGVPG 733
Db 588 SGLSHLSHLDIYHDIKPNQNLISLPDAKGVVMISDFGLCKLNFQKTSFSRSRGVPG 647
OY 734 TEGMIAPEML-----SEDCKENPTY-----TVDFISAGCYFTYV 767
Db 648 TDGWIAPPEMRSQRTVRHDFYUYPILLHNNKFLKSMSEFNLYQTTAVDIFISGCVYVV 707
OY 768 VEGSHPEGKSLQROANITLIGACSLDCLHPEKHED---VIARELIEKMIADPOKRPQA 823
Db 708 LSGHHHAFGDNLRQANILISHEYNLAKLREDDSEDSRIILAEOLISDHIKQPOSRPA 767
OY 824 NDVLKHPFMSLEKOLQFQDVSDRIEKSLLDGPVYKOLERGRAVVKMDMRENTIDPQ 883
Db 768 RIGNHPLFEMWDEPKMLSFQDVSDEVEKQFHAERPLKSEKNGRIYVLDWMVHIDPMIT 827
OY 884 TDLKRFRTYKGSVDLRLAMRNKKHNYRELPAEYARETIGTLPDQFVCTYSRFPYLLAH 943
Db 828 DDLRYRGYMGASVDLRLALRNKKHNYHELTPAAOKMGLGCIPIHETNTVWVDFPOLISH 887
OY 944 TYRAMELCSHERLQPYU-----FHHP-----PEQOPRV 972
Db 888 AYHAFSISGNEPIFKPYISAGYLFTRPMYFDADDALEPMLMDPKLPKPI 937
```

```
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PROBABLE SERINE/THREONINE-PROTEIN KINASE C167.01 PRECURSOR
GN SPAC167.01.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Rieger M., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (JAN-1999) to the EMBL/Genbank/DBS databases.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (BY SIMILARITY).
CC - PTM: AUTOPHOSPHORYLATED PRIMARILY ON SERINE RESIDUES (BY
CC SIMILARITY).
DR EMBL: AL035248; CAA22846.1; -.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR InterPro: IPR002290; Euk_pkinase.
DR Pfam: PF01011; Bacterial_PQO; 2.
DR Pfam: PF00069; pkinase; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; Transmembrane; Signal; Transferase;
KW Serine/threonine-protein kinase; ATP-binding; Glycoprotein;
KW Endoplasmic reticulum; Phosphorylation.
FT SIGNAL 1 30
FT CHAIN 31 1072
FT FT PROBABLE SERINE/THREONINE-PROTEIN KINASE
FT C167.01
FT FT
FT DOMAIN 31 518 LUMENAL (POTENTIAL).
FT TRANSMEM 519 539 POTENTIAL.
FT DOMAIN 540 1072 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 29 34 POLY-SER.
FT DOMAIN 654 938 PROTEIN KINASE.
FT NP_BIND 660 668 ATP (BY SIMILARITY).
FT BINDING 682 682 ATP (BY SIMILARITY).
FT ACT_SITE 774 774 BY SIMILARITY.
FT MOD_RES 813 813 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 814 814 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 1072 AA; 121222 MW; 5B1A60EEFC802526 CRC64;

Query Match 18.7%; Score 959.5; DB 3; Length 1072;
Best Local Similarity 27.4%; Pred. No. 1,5e-64;
Matches 296; Conservative 167; Mismatches 367; Indels 249; Gaps 36;

OY 35 VETLDGSLAVSKRTGSIKWTL-----KEDPVLYQ-----PTNVE- 69
Db 87 VATVDGSLSYDRITQGLMSLFTNANPGLSTYTKDENSLSKFSQNFKYSNTHGEF 146
OY 70 -----EPAPLPDPDGS-LYTLGSKNNEGJTKLPFTIPELVQASPCRSSDG 114
Db 147 YSDSTLINISYSDDDTVWVEVPEIDGILYAFNLQF--GLVRLPHSFKIDLVHSPRLNN 204
OY 115 ILXMGKKQDIWYIDLTGE-----KQQLSSAFADS----- 146
Db 205 NNFVSGSKNTLFTFIDVNSGDIVSQYPSGHRYTEHSHVNLGKRRDPSVPSG-ADSDLSFSD 263
OY 147 -----LCPSTSL-----YIGREYVITVYDTRKTRLRNNATYFDYAAAL 186
Db 264 PSCKKLSESLDLDPENYQVTVSNKSFVIAETRTYITYSNSNTL--DLVYIDMTPT- 320
OY 187 PEDEGDYKMSH-----FVSGNGDLVYTVDSGVDLMIQNYASPVAAFYVMQREGIR 238
Db 321 -KNEIMYSEFSSPDSKALASSYDSLHIVDHSFKIR--QNIPLMSPAAYFDIVLTP 377
OY 239 KYMHINAVETLRILTFMSEGEVGRITKMKYPPPKETEAASKLPTLYVVKYSTSLAASP 298
Db 378 HNKIKIDSQTPAKFPPTSVLLRQPIDTYLETMPQIARNK--TEHYVNHIGNAMFAMSE 434
```

```

OY 299 MHEGAVVPRGSTL-----PLLEGPODGVIT---GDKE 331
DB 435 RHPPLVSLAEASFLYNGFYPLNLSIFGLSLMAYPKPFALPGJGYPVSESGS 494
OY 332 CVITPNDVY----FDGKSKNKLNLYRWYLLIGHHEPRLASTKMLERFPNNLPKH 367
DB 495 TKLPFGKKPRLDNPISSTPIST--TFWIMF-----LVSFTIYVF--SILKIRS 546
OY 388 ENVIPADSEKSEFEVYINLVDOJTSNAPTVSRDVEEKPAHAPAPAPVDSMLKMATI 447
DB 547 SEVRPLKSGQNTV--SINKIDTSKRRKKRKRKRVSE--HSAS-----587
OY 448 ILSTFLICGVAFITTYPLSMHQQOOLQHOQFKELEKLOLLOQOQOOLPFHPRGTAOD 507
DB 588 --SNF-----NELESQASEFQONTDILESENIEI-----QD 617
OY 508 GELDQSGPYSESSGTSPTSPRASNHLSCSSSASKAGSPSLBDDODEFTSVYIG 567
DB 618 ----KTDPLQKSLDSSLSHLPEAT-----VIONTDS-----VTYN 651
OY 568 KISCFKPDVLGHGAEGLIYVGMFNDRVAVKRIPECFSEFADREVOLLRESDENPVIR 627
DB 652 SLTVVP--EVIGYSGHGIYVRYGYEDREVAVKRVLMFVYOLASREITLLQSDNHPVIR 710
OY 628 YFTEKDRPOFYAIELCAATLOEYVEQK---DPAHLGLEPTLLQOTTSGLAHLHSLNT 684
DB 711 YKCKQSDFLYIYIELCKNLSDLEIKPLAYDLPKSIDVLVLYOIAVYSHLSL 770
OY 685 VHHDLPHNLISMNPNHG---KIKAMISDFGLCKKLANGRHSFGRS--GVPTGEMIA 739
DB 771 VHHDLPHNLISLVNNSPNLSKTRRALISDFGLSKKLDPNOSLSKNTTTEAAGSTGWS 830
OY 740 PEM-----LSECKE-----NPTYVDIFSAGCYEYVVSSESHPFKSLQROAN 784
DB 831 PELSSLSQSGSEIQVKTREGRIRQASHATDIFALGCIYVTLTGSMHFGSHYCEGN 890
OY 785 ILGACSLDCLNPKHEEDYIARLEKMTAMDQKPSANDVUKHPFMSLEKOLQOFOD 844
DB 891 ILGNANCLVHLQSLGEGVLAADLIEDIMAFEPSPRTIEVVLNHPFLFMDYAKKLDLID 950
OY 845 VSDRIKESLD--GPIYKOLERGSAVVVKMDRENTDPLQDTLRKRYTKGGSVBDLR 902
DB 951 VSDREVEERDPPSLQMLNNSKSVIGENMTTCILHSLVDNLGKTRKIDGSKIDILR 1010
OY 903 AMRNKHHYRELPAEYRETLGLTLPDFVCYFTSRPHLLAHTYRAM--ELCSHERLFQY 960
DB 1011 VLNRKHHHQDLPESEVRVVLGDLPGFTSYFEKFPMLLHCHLVKDVLYEESQFKRY 1069

RESULT 7
OYFS17 PRELIMINARY; PRT; 893 AA.
AC OYFS17 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE OSIREL.
GN OSIREL.
OS Oryza sativa (Rice).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Okushima Y., Koizumi N., Sano H.;
RT Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB031396; BAB20385.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR InterPro: IPR001245; Tyr_kin.

```

```

DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PRO0109; TYRKINASE.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYRKc; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 893 AA; 99087 MW; E32EC6BFC5161255 CRC64;

Query Match 15.0%; Score 770; DB 10; Length 893;
Best Local Similarity 28.6%; Pred. No. 4e-50; Indels 158; Gaps 30;
Matches 234; Conservative 126; Mismatches 300;

OY 242 HINVAETTLFTFMG-----EV-----GRTKKY-----PPKET--EAKSKLT 281
DB 131 HITSKPDNLVYIYLDGSEHSDILEVHNGSVALLPKLEFIAETPTIYINDSFYTISKS 190
OY 282 PLYVCKYSTSLYASPMHEGAVVPRGSTLPLE--GPQIDGVITGDKGECVITPSTD 339
DB 191 TTFVNVADSGEIIYKHSLS--PVALNEVG--PLVEIISKLDARSSTSANIIVVVRD 245
OY 340 VKFDPGLKSKNKLNLYR---NYWLLIGHHEPRLASTKMLERFPNNLPKHRENVIPAD 394
DB 246 YTSIASDGLGELHLMWTRTSFTANYTARYGHQD--MLAQSSCLR--GNIPCTIRGEPYK 300
OY 395 SEKSEFEVENVLDQTSNAPTVSRDVEEKPAHAPAPAP-----VDSK 440
DB 301 -----LYLPDSSDNA--IVLRPVNEVSAVDALPELPPKKLPAPAGESNVALDSA 349
OY 441 LKMATIILS-----TFLICGVAFITTYPLSMHQQOOLQHO 477
DB 350 QNOADIALGHFVADTELNTSYTKESYRWLPFTFLMLIMACLVKLADASKYCRQF--- 406
OY 478 QFQKELEKLOLLOQOQOOLPFHPRGTAODGELDQSGPYSESSGSPSTSPR----- 531
DB 407 -----VIRELK-----PF-----MRDEKIMD--PRKSGESTSKRRKAKKQDGLN 444
OY 532 ASNHSLSGSSSASKAGSPSLBDDODEFTSVI-----VGKISCFKPDVLGHGAG 583
DB 445 STQIFASDKEGNGTGGSTPAQSNKANDSTINVELPGLNGRQGLKCYVSKEL--GAGSNG 503
OY 584 TTYVGMFNDRVAVKRIPECFSEFADREVOLLRESDENHPVIRYCTEKRQFOYIAIE 643
DB 504 TYVEEGSYGGEVAVRRLRSHNDIASKEIENLADODPNIVRMYGFEQDNDFFVYISLE 563
OY 644 LCAATLOEYVEQK---FAHL-----GL-----EPITLLQOTTSGLAHLHSLNIYH 686
DB 564 KCRCSLADLTQLHVSPPFSNTKGTIDELMRQDLPQAOLIKLMRDVAVAGIVLHSLGITH 623
OY 687 RDLKPNHILISMPNAGKIKAMISDFGLCKKLAVGRHSFSRSGVPGTEGWIAPEMLSFD 746
DB 624 RDLKPNVILS---KSGPLRAKLSDMGISKRLQEDMTSVSHGTGSGSGWQAPEDLRNG 680
OY 747 CKENPTYVDISGACVFIYVVSSESHPRGKSLQROANTILGACSLDCH--PEKIEDVI 804
DB 681 ---RQTRADLDFSLGLIFCITKRGKHPGEYERDMKIIINOFDILFIYDHIPE----- 731
OY 805 ARELIEKMTAMDQKPSANDVUKHPFMSLEKOLQOFQDVSRIKESLDGPIYKOLR 864
DB 732 AVHLISQLDLPDEKRPATVYVNHHPFMSPELCSFLKDTSDRIKTS--ETDILALAG 790
OY 865 GGRAVVKMDRENITDPLQTLRKFTYKGSVYDOLLAMRNKHHYRELPAEYRETLGT 924
DB 791 INVEAFKWMGKELDAALLADMGRYKYSPESTRDLRLIRKNSGYHREFSDLKLLOS 850
OY 925 LPDDEVCIYTSRPHLLAHTYRAM--ELCSHERLFQY 961
DB 851 LPEGVQVYESSRPPKLLIRVYEMSEHCKDEAFSKYF 888

RESULT 8
OPENING

```